SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Context-Aware Representation Learning in High-Dimensional Dynamic Systems

Ezhilarasan Ganesan[1], Pushpa Nagini Sripada[2], T. Shanthi[3], Arvind Kumar Pandey[4], Mukesh Patidar[5], Pushpalatha P[6], M. N. Vimal Kumar[7]

[1] Professor, Department of Electrical and Electronics Engineering, Faculty of Engineering and Technology, JAIN (Deemed-to-be University), Bengaluru, Karnataka, India. Email: g.ezhilarasan@jainuniversity.ac.in, ORCID: 0000-0002-5335-2347
[2] Professor, English, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, India. Email: sripadapn@maher.ac.in
[3] Associate Professor, Department of Electronics and Communication Engineering, Sona College of Technology, India. Email: shanthi@sonatech.ac.in, ORCID: 0000-0001-8962-5311
[4] Associate Professor, Department of Computer Science & IT, Arka Jain University, Jamshedpur, Jharkhand, India. Email: dr.arvind@arkajainuniversity.ac.in, ORCID: 0000-0001-5294-0190
[5] Associate Professor, Department of Computer Science and Engineering, Faculty of Engineering and Technology, Parul Institute of Technology, Parul University, Vadodara, Gujarat, India. Email: mukesh.patidar34885@paruluniversity.ac.in, ORCID: 0000-0002-4401-8777
[6] Assistant Professor, Computer Science, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, India. Email: pushpalathap@maher.ac.in
[7] Department of Mechatronics Engineering, Sona College of Technology, Salem, Tamil Nadu, India. Email: vimalkumar.mct@sonatech.ac.in

Abstract

Rapid advancements in remote sensing technologies and distributed monitoring systems have enabled the large-scale acquisition of high-dimensional geo-spatiotemporal data in agricultural and environmental domains. However, extracting meaningful representations remains challenging due to complex spatial–temporal dependencies and limited labeled data availability. To develop an accurate and robust framework for crop yield prediction by learning context-aware representations from high-dimensional dynamic agricultural data. Multi-source datasets are utilized, including satellite imagery, ground-based sensor measurements, and historical environmental records. Key variables include vegetation indices (NDVI, EVI), temperature, rainfall, soil moisture, humidity, solar radiation, and temporal growth indicators. Data preprocessing involves Z-score normalization to standardize feature distributions, followed by Principal Component Analysis (PCA) for dimensionality reduction and feature extraction. The proposed model integrates Red Panda Optimized Variational Autoencoder with attention mechanism (RPO-VAE-ATT) to jointly capture spatial structures and long-range temporal dependencies. The proposed RPO-VAE-ATT model uses VAE for extracting compact representations from high-dimensional data, RPO for optimizing parameters, and attention to capture temporal dependencies, enabling accurate learning of spatial–temporal patterns for robust and efficient crop yield prediction. The framework achieved high predictive performance with improved accuracy (95.3%), precision (95.03%), recall (96%), and F1-score (97.39%), demonstrating strong generalization and robustness under data-scarce conditions, and was implemented using Python tools. The framework provides efficient and adaptive representation learning for high-dimensional dynamic systems, enabling reliable crop yield prediction and supporting data-driven decision-making in precision agriculture.

Keyword: Context-Aware Representation Learning, Geo-Spatiotemporal Data, Temporal Dependency Modeling, Remote Sensing, Crop Yield Prediction, Deep Learning

Introduction

Multidimensional data of the complex environment is still growing exponentially in the contemporary world. As a result, creating precise and effective models for high-dimensional dynamic systems has started to pique the interest of researchers and practitioners [1, 2]. Complex systems, like those observed in finance, healthcare, engineering, and smart technologies, typically comprise many factors that interact over time in a complex and often non-linear way [3]. These dynamic systems are high dimensional, dynamic and highly difficult to model satisfactorily by traditional modelling methods. By enabling models to produce meaningful and adaptive representations of dynamic data, Context-Aware Representation Learning (CARL) has recently surfaced as a novel paradigm for solving some of these challenges related to modelling high-dimensional dynamic systems [4]. Furthermore, by collecting intricate spatiotemporal patterns, integrating environmental data from several sources, and facilitating adaptive, data-driven decision-making for precision and sustainable agricultural techniques, CARL helps agriculture anticipate crop yields accurately [5].

CARL combines contextual and temporal data to generate a more comprehensive picture of underlying patterns and relationships existing in complex environments, unlike more traditional representation-learning methods, which only use contextual data as training or testing models [6]. The effective modelling of both the static and dynamic properties of high-dimensional data with CARL contributes greatly to predictive performance, robustness and generalisation of the generated models [7]. Moreover, CARL allows being more flexible in adapting to the changes in a specific environment due to the ever-changing high-dimensional data [8]. Finally, although equally importantly, reduce noise and duplication in their high-dimensional data by employing CARL. Finally, although equally importantly, reduce noise and duplication in their high-dimensional data by employing CARL. Consequently, they are more capable of extracting only the features that are necessary for downstream activities such as decision-making, anomaly detection [9, 10]. The key research objectives are as follows,

- To collect and preprocess multi-channel geo-spatiotemporal data from satellite imagery, sensor networks, and historical records, and extract meaningful features representing dynamic environmental variations.
- To design a VPO-VAE-ATT model capable of capturing spatial structures, modelling long-range temporal dependencies, and enhancing global contextual learning through a learnable query-based attention mechanism.
- To improve prediction accuracy and adaptability by developing a scalable and efficient framework for modelling high-dimensional dynamic systems.

The research is organized as follows: Section 1 introduces the background of the research, Section 2 reviews related work, Section 3 presents the proposed VPO-VAE-ATT methodology, Section 4 discusses results and evaluation, and Section 5 concludes with findings and future work.

Review of Literature

This section reviews existing Machine Learning (ML) and Deep Learning (DL) models was crop yield estimate, remote sensing-based organization, and precision agriculture, summarizing their performance and limitations

DL for E-Commerce Prediction and Decision Intelligence

To utilise Natural Language Processing (NLP) with ML and DL techniques to create a context-aware Structured Query Language (SQL) injection detection system [11]. The DL model was highly effective as the accuracy was high with low error rates (0.14%). Nevertheless, the fact that the approach relied on datasets turned out as a drawback, and it requires further testing in various real-life scenarios. This study was on prediction, personalisation, and decision intelligence as it explored newer advances in DL in the context of e-commerce [12]. Using AI and reinforcement learning approaches, it examined applications like fraud detection, demand forecasting, recommendation systems, and operational optimisation. The findings showed that DL was able to enhance the functioning of large-scale platforms and decision making. Scalability, robustness, interpretability, and cross-border adaptation, however, continued to be significant constraints. They suggested a hybrid intrusion detection system to detect high-dimensional data, class imbalance, and dynamic traffic in IoT-cloud environments [13]. To identify and categorize abnormalities, it integrated VAE, Isolation Forest, and Graph Attention Network (GAN). It outperformed the existing systems in terms of performance. However, reliance on dataset attributes and additional validation in an evolving actual setting were the drawbacks. A contextual sarcastic recognition model was developed using user profiles and forum topic data [14]. Semantic and contextual characteristics were extracted using a fusion network and a Bidirectional Long Short-Term Memory (Bi-LSTM) with spatial attention. The results showed improved performance over existing methods. Limitations were that it used contextual data and could not be easily generalized across platform.

ML and DL Approaches in Prediction, Detection, and Crop Yield Analysis

In Pakistan, remote sensing indices and climatic variables have been used to predict the yield of wheat [15]. Applied Random Forest (RF), Support Vector Machine (SVM), Least Absolute Shrinkage and Selection Operator (LASSO) and Ridge models under two seasonal conditions with various combinations of features. RF did the best (Highest R^2 to 0.78, smallest Mean Absolute Error (MAE) and Root Mean Square Error (RMSE)). Assessed an ML framework based on phenology winter wheat sub-field yield prediction [16]. Dl techniques were employed. RF was the best performer (R^2 0.73, RMSE 0.88 t/ha) and accuracy was better with joint phenological and climatic features. To denoted minimal utility of soil data at the scale being investigated and spatial error. Created feature-Enhanced Class-Attention Residual TransUNet (FCR-TransUNet) to classify crops using remote sensing properly [17]. TransUNet with Feature Enhancement Module and Class-Attention to minimize noise, class confusion and blurred boundaries. Mean Intersection over Union (MiosU) of 92.2% (Youyi Farm) and 89.9% (barley), exceeding those of the baseline models, and being stable across growth stages. Developed a framework for anomaly detection and resource optimization in precision agriculture [18]. Integrated Multi-Modal Smart Farming Network (IMSFNet) with the Adaptive Resource Optimization Strategy (AROS) using multimodal data and spatiotemporal modeling. Improved anomaly detection accuracy and decision-making efficiency. Required high computational power and large, high-quality datasets. Developed a fuzzy hybrid ensemble method for crop classification and yield estimation using remote sensing data [19]. Applied image enhancement, data augmentation, and bagging-based ensemble learning on multisource agricultural datasets. Achieved higher classification accuracy and lower prediction error compared to traditional and individual models. An effective hybrid DL model that predicts rice crop yield by coordinating the results of a regression model with a classification model in DL through shared layers [20]. DL methods include ensemble-driven learning techniques, Artificial Neural Networks (ANNs), and Multiple-Parametric Deep Neural Networks (MDNNs). Outperformed current models in terms of accuracy (90%) and F1 score (94%).

Research gap: Existing crop yield prediction models often rely on large datasets, limiting performance under data-scarce conditions [15]. Most approaches use single-modal data, ignoring multi-source spatiotemporal integration [19], and many require high computational resources and extensive preprocessing [20]. To address these gaps, the framework integrates multi-source geo-spatiotemporal data with context-aware representation learning. A hybrid VAE with ATT captures spatial and temporal dependencies. Semi-supervised learning leverages unlabeled data to improve robustness under limited labeled conditions.

Methodology

Develop context-aware representations from high-dimensional dynamic agricultural data to create a reliable and accurate framework for crop yield prediction. Multi-channel data are collected from satellite imagery, sensor networks, and historical environmental records. That proposed VPO-VAE-ATT model integrates an optimized VAE with an attention mechanism to capture spatial and temporal dependencies. A semi-supervised learning approach enhances robustness and improves prediction performance, as shown in Figure1.

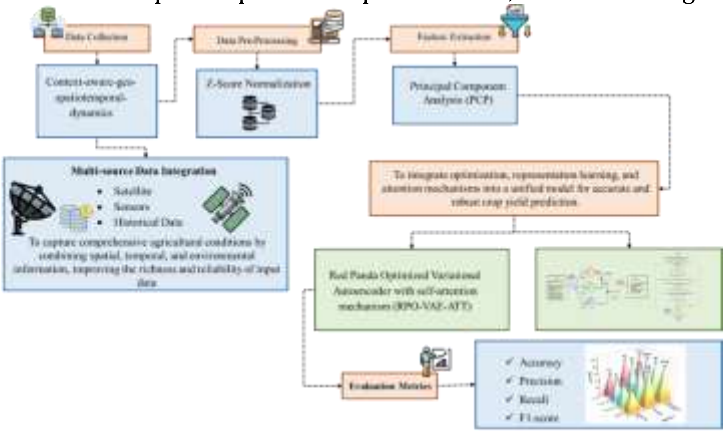

Figure 1: Proposed Framework for Context-Aware Crop Yield Prediction Using RPO-VAE-ATT Model

3.1 Data Description

Context-Aware Geospatial and Temporal Dynamics this Kaggle dataset is high-dimensional, multi-source information that describes dynamic environmental and agricultural events over time and space. Its 9,201 rows and 15 columns include satellite images, data recorded on the ground by sensors and historical records. Important factors include temperature, rainfall, soil moisture, humidity, solar radiation, vegetation indices (NDVI, EVI) and temporal growth indicators. To provide reliable context-aware representation learning and precise crop yield prediction in intricate dynamic systems, the dataset were preprocessed and divided into 70% training and 80% testing sets.

Source: https://www.kaggle.com/datasets/programmer3/context-aware-geo-spatiotemporal-dynamics-dataset/data

Data preprocessing using Z-Score Normalization

To improve the accuracy of crop yield forecast by standardized, cleansing, and normalizing data from agriculture for dependable and consistent model inputs. Numerical variables are scaled using Z-score algorithm to bring the features up to a mean of 0 and a standard deviation of 1 to achieve consistency and enhance the performance of the model. The contribution made by these variables to improving analytical capability and predictive accuracy is shown in Equation (1).

$$D_{new} = \frac{D-\mu}{\sigma} = \frac{D-\mathrm{mean}(D)}{\mathrm{StdDev}(D)} \quad (1)$$

D represents the original data or feature value, while D_{new} denotes the normalized value obtained after applying Z-score standardization. μ is the mean of the original data $D-\text{mean}(D)$, σ is the standard deviation of the original data StdDev(D). The original data's average is denoted by the term mean(D), while the data's standard deviation is denoted by the phrase StdDev(D).

Principal Component Analysis (PCA): Reduces dimensionality by extracting key informative features efficiently

Through reducing multidimensional agricultural data into important features, removing redundancy, and improving computing power while maintaining crucial information influencing crop yield production. PCA is used as a complexity reducing method to accomplish this, reducing multidimensional data to a smaller number of unrelated elements known as primary elements. The projection of normalized data onto principal components is given by Equation (2-8),

$$E_j = V_i^T N = \sum_{j=0}^{n} V_{ij} N_j \tag{2}$$

E_j denotes the projected feature value, V_i^T is the transpose of the eigenvector, V_{ij} represents the eigenvector element corresponding to the j^{th} feature, N is the normalized input vector, N_j is the j^{th} normalized feature, and n is the total number of features represented in Equation (3).

$$\text{Var}(N) = \frac{1}{m}(VN)(VN)^T = \frac{1}{m}VNN^TV \tag{3}$$

Var(N) is a variance of normalized data, $\frac{1}{m}$ is several samples**,** VN is a eigenvector matrix,$(VN)^T$ is a transpose of eigenvector matrix, N is a normalized data matrix, N^T is a transpose of normalized data matrix, VN is a projection of data onto principal component space, VNN^TV is a covariance structure of normalized data depicted in Equation (4).

$$\text{MaxVar}(N) = \text{Max}\left(\frac{1}{m}VNN^TV\right) \tag{4}$$

MaxVar(N) denotes the maximized variance of normalized dataset features, V represents the eigenvector matrix selecting dominant structural patterns, N indicates normalized feature inputs, and m is the number of principal components are presented in Equation (5).

$$\frac{1}{m}VNN^T = \text{cov}(N),\ \text{Var}(N) \tag{5}$$

$\frac{1}{m}VNN^T$ represents the scaled covariance structure of normalized dataset features, cov(N) denotes the covariance distribution of the dataset, Var(N) indicates variance capturing dominant patterns, and m is the number of selected principal components (Equation 6).

$$\text{Var}(N) = V^T\text{cov}(N)V \tag{6}$$

Var(N) is the variance of projected data along the principal component, V is an eigenvector, V^T is the transpose of the eigenvector, cov(N) is a covariance matrix of normalized data, and N is a normalized dataset (Equation 7).

$$K = V^T\text{cov}(N)V - \delta(V^TV - 1) \tag{7}$$

K is an objective function in PCA, VEigenvector, V^T is a transpose of eigenvector,cov(N) is a covariance matrix of normalized data, N is a normalized dataset,δ is a Lagrange multiplier, V^TV is a norm of eigenvector, and 1 is a unit constraint ensuring eigenvector has length 1.

$$\frac{dK}{dV} = 0 \Rightarrow \text{cov}(N)V - \delta V = (\text{cov}(N) - \delta J)V = 0 \tag{8}$$

In Equation (8), K is a variance captured by principal components, V is at eigenvector, $\frac{dK}{dV}$ is a optimization condition to maximize variance, N is an input dataset matrix, cov(N) is a covariance matrix of dataset N, representing feature relationships, δ is an eigenvalue. J is the identity matrix, and PCA maximizes variance to select optimal components, supporting improved representation learning and prediction performance.

Red Panda Optimized Variational Autoencoder with attention mechanism (RPO-VAE-ATT): Learns context-aware representations capturing spatial-temporal dependencies for accurate crop yield prediction

The proposed model, RPO-VAE-ATT, is an effective approach for modeling high-dimensional dynamic agricultural data and improving crop yield prediction, as it integrates optimization, deep representation learning, and attention mechanisms. VAE learns compact latent representations from multi-source inputs, capturing nonlinear patterns, while attention enhances context awareness, and RPO optimizes hyperparameters for faster convergence and improved accuracy.

Variational Autoencoder (VAE): Learns latent representations capturing complex distributions from high-dimensional input data.

To assist in the prediction of crop yield production by using complicated agricultural data to create compact latent representations that capture invisible trends and variation that affect crop growth and yield outputs. To enhance latent feature learning and interpretability in high-dimensional dynamic systems using an improved VAE. The proposed method maps attention outputs into the latent space and employs the reparameterization trick for stable optimization.. The Evidence Lower Bound (ELBO) is defined as Equation (9),

$$\mathrm{ELBO} = \mathbb{E}_{r_\phi(y|W)}[\log \rho_\theta(W \mid y)] - \mathrm{KL}\big(r_\phi(y \mid W) \parallel \rho(y)\big) \tag{9}$$

The input W represents high-dimensional system data, and $y$is the latent variable encoding key features. The encoder $r_\phi(y \mid W)$ maps $W$to a distribution over y, while the decoder $\rho_\theta(W \mid y)$reconstructs W from y. KL divergence $KL(r_\phi(y \mid W) \parallel \rho(y))$measures deviation from the prior $\rho(y)$, and ELBO maximizes reconstruction accuracy while regularizing the latent space. A modified ELBO with a balancing parameter is expressed as Equation (10).

$$\mathrm{ELBO} = \mathbb{E}_{r_\phi(y|w)}[\log \rho_\theta(w \mid y)] - \beta\,\mathrm{KL}\big(r_\phi(y \mid w) \parallel \rho(y)\big) \tag{10}$$

W is the input data, y is the latent variable, $r_\phi(y \mid W)$is the encoder mapping inputs to latent distributions, $\rho_\theta(W \mid y)$ is the decoder reconstructing inputs from latent space, $\rho(y)$ is the prior distribution of y, $KL(\cdot)$ measures divergence between encoder and prior, βbalances reconstruction and latent regularization, and $\mathbb{E}_{r_\phi(y|W)}[\cdot]$ represents the expected reconstruction likelihood. The encoded latent representation is obtained as Equation (11).

$$\hat{w}_s = e_\theta\big(Y_{(s-S:s)}\big) \tag{11}$$

$\hat{w}_s$denotes the learned latent representation at step s, $e_\theta(\cdot)$is the encoder function, and $Y_{(s-S:s)}$represents the input sequence over a temporal window of size S. The conditional output distribution is defined as Equation (12).

$$O_\theta\big(w_s^j \big| Y_{s-S:s}\big) = M\big(w_s^j \big| w_s^j, \sigma_j^2\big) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{\big(w_s^j - \hat{w}_s^j\big)^2}{2\sigma_j^2}\right) \tag{12}$$

Where $\big(w_s^j \big| Y_{s-S:s}\big)$ is the observed value of feature $j$at time s, $\hat{w}_s^j$is the predicted mean, and σ_j^2represents prediction uncertainty. $Y_{(s-S:s)}$denotes past observations, and $\theta$are the model parameters. The Gaussian $M(w_s^j \mid \hat{w}_s^j, \sigma_j^2)$computes the likelihood of w_s^j, enabling robust modeling of temporal sequences. This enables compact, context-aware representations that improve prediction accuracy in complex dynamic systems. The attention weight is computed using Softmax as Equation (13),

$$b_s^l = \frac{\exp\,(f_s^l)}{\sum_{j=1}^{m} \exp\,(f_s^j)} \tag{13}$$

b_s^l is a normalized probability of feature l at time step s. f_s^l is a raw score lat time s. f_s^j is a raw score of all classes $j = 1,2,\ldots,m$ at time s. m is a total number of features being considered. exp $(\cdot)$– is an exponential function applied to logits to ensure positivity. $\sum_{j=1}^{m} \exp\,(f_s^j)$ is a normalization term, summing exponentials over all classes to ensure probabilities sum to 1 given by Equation (14).

$$\bar{y}_s = \sum_{l=1}^{S} b_s^l\, g_l \tag{14}$$

$\bar{y}_s$ is the context-aware representation, b_s^l are attention weights, g_l are hidden states, and S is the sequence length

Attention mechanism (ATT): Focuses on important spatial-temporal features enhancing context-aware learning capability.

Through identifying important growth stages and capturing temporal connections, context-aware training from dynamical agricultural and environmental data can enhance crop yield prediction. Attention mechanism allows the model to flexibly give importance to various spatial and temporal regions, in contrast to typical convolutional procedures that concentrate on local neighborhoods. This is especially crucial when predicting agricultural yields because the accuracy of predictions is greatly impacted by interactions between distant features, such as environmental conditions over time and distance. The model learns more context-aware and discriminative feature representations by incorporating attention into the RPO-VAE framework. Using learnable weight matrices and the 1×11 times 11×1 convolution operations provided by Equation (15), this feature map is converted into three distinct representations.

$$f(x) = W_f x,\ g_x = W_g x \tag{15}$$

The input attribute map is denoted by x. The input is projected into $f(x)$ and $g(x)$, two distinct feature spaces, using the learnable weight variable matrices W_f and W_g. In essence, $f(x)$ and g_x assist the model grasp one portion of the data connects to another region, which forms the basis for attention computation. The proximity between different points in the feature map provided by Equation (16 & 17) is determined using these altered properties.

$$B_{i,j} = \frac{exp\ (f(x_i)^T g(x_j))}{\sum_{i=1}^{N} \exp\ (f(x_i)^T g(x_j))} \tag{16}$$

The variable x represents the input mapping of features obtained from farm data, and the characteristic vectors at places i and j are indicated by the variablesx_i andx_j. Adaptable weight matrices are used to alter the depiction of function $f(x_i)$ and $g(x_j)$. The similarity of features is measured by their dot product, $f(x_i)^T\ g(x_j)$. The emphasis weight indicating the $i-th$ attribute's impact to the j-th output is denoted by the term $B_{i,j}$. In this case, $\sum$ normalizes the scores into a spectrum of probabilities $exp(\cdot)$ guarantees positive values, and N is the sum of feature values. Equation (17) provides the final output after the attention output is scaled using a reusable parameter and coupled with the initial input feature mapping:

$$y_i = \gamma o_i + x_i \tag{17}$$

The attention layer's ultimate output is represented by Equation (17). The original input feature is denoted by x_i while the attention-enhanced feature at position i is denoted by the word o_i. The scalar parameter γ regulates the attention mechanism's input. The model successfully captures both local component details and global situational constraints by including this attention mechanism into the RPO-VAE framework. Figure 2 illustrates how a multi-branch architecture reassembles outputs for robust context-aware representational learning after extracting hierarchical features and integrating them via attention-based fusion.

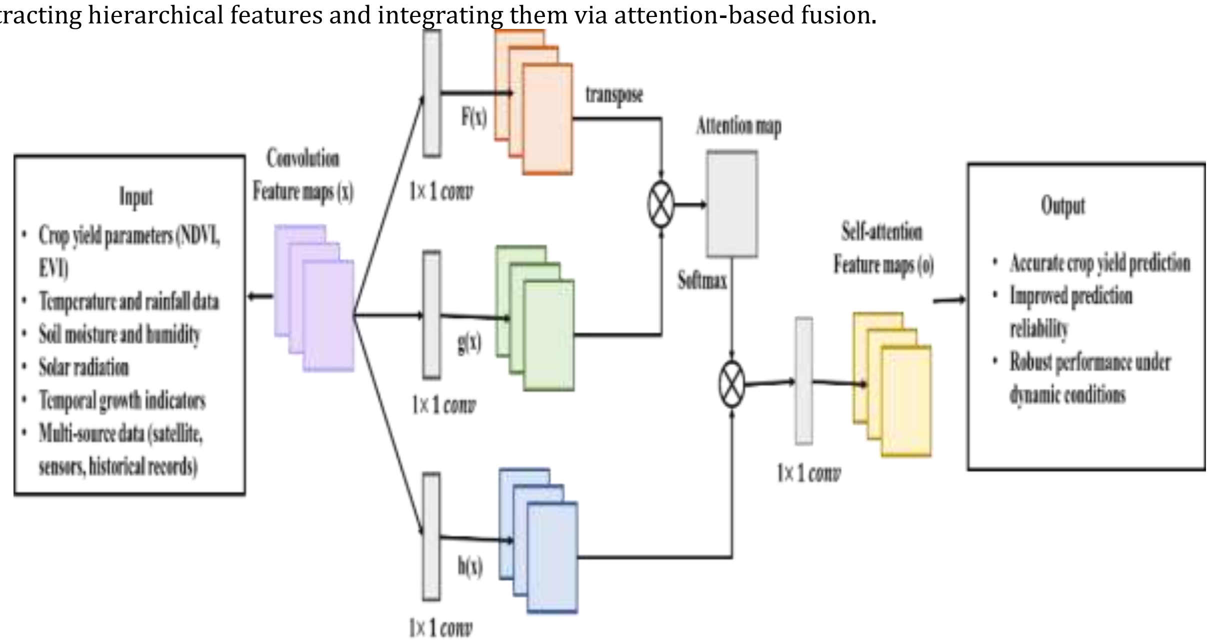

Figure 2: Multi-Branch Deep Feature Fusion Architecture with Attention Mechanism

RPO: Optimizes hyperparameters efficiently to improve model convergence and predictive performance.

Optimizing model parameters, enhancing convergence, lowering prediction errors, and guaranteeing effective learning for farming data trends are all ways to improve crop yield prediction accuracy. The best candidate solution represents the optimal learned representation, ensuring effective tuning of VAE-ATT parameters and enabling accurate, context-aware, and scalable representation learning for high-dimensional geo-spatiotemporal data as shown in Figure 3.

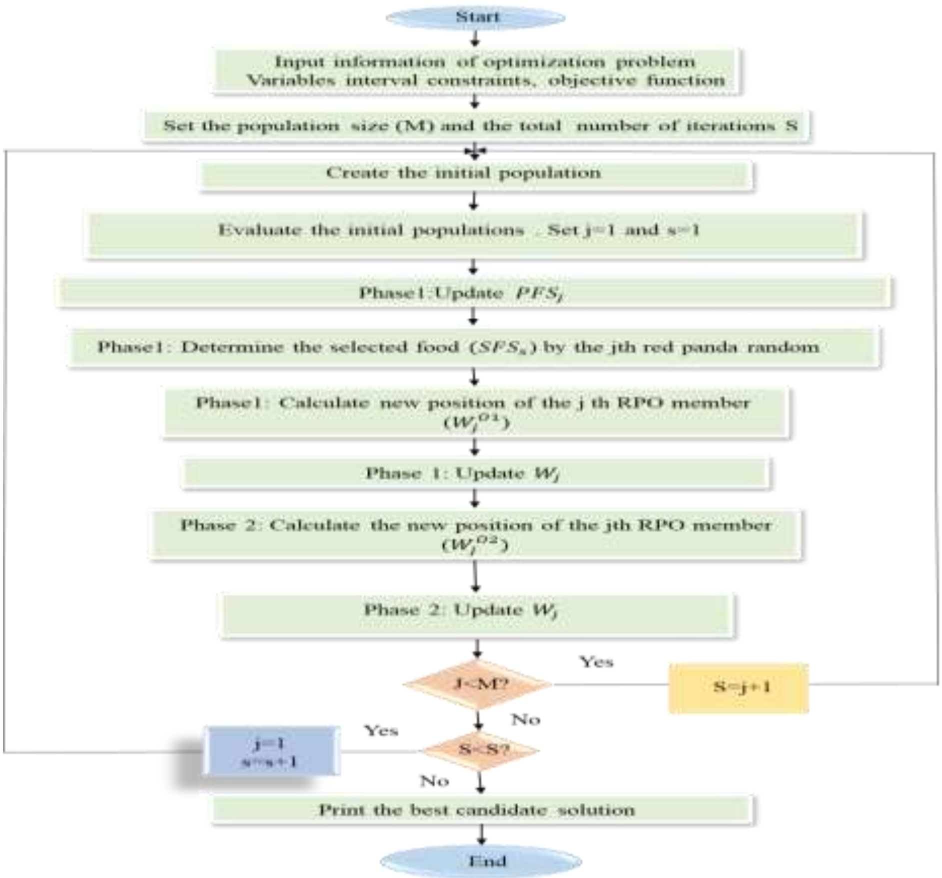

Figure 3: Flowchart illustrating the RPO algorithm process.

Phase 1: Foraging Strategy (Exploration): Enhances global search to capture complex spatial-temporal dependencies and avoid poor local representations. The promising foraging set is defined as Equation (16).

$$\mathrm{PFS_j} = \{\mathrm{W_l} \mid \mathrm{l} \in \{1,2,\ldots,\mathrm{M}\}, \mathrm{E_l} < \mathrm{E_j}\} \cup \{\mathrm{W_{best}}\} \tag{16}$$

$\mathrm{PFS_j}$ denotes the promising solution set for the $\mathrm{j^{th}}$candidate, $\mathrm{W_l}$represents candidate solutions, M is the total number of candidates, $\mathrm{E_l}$and $\mathrm{E_j}$are fitness values, and $\mathrm{W_{best}}$is the best solution. The position update is given by Equation (17).

$$\mathrm{W_{j,i}^{O1}} = \mathrm{w_{j,i}} + \mathrm{q} \cdot (\mathrm{SFS_{j,i}} - \mathrm{J} \cdot \mathrm{w_{j,i}}) \tag{17}$$

$\mathrm{W_{j,i}^{O1}}$ is the updated value of the $\mathrm{i^{th}}$ dimension of the $\mathrm{j^{th}}$ candidate, $\mathrm{w_{j,i}}$ is the current value, q is a random coefficient controlling exploration, $\mathrm{SFS_{j,i}}$ is the selected foraging solution, Jis a scaling factor, and i, j denote dimension and candidate indices.

Phase 2: Climbing and Resting (Exploitation): Refines latent representations locally, improving feature consistency and model generalization in dynamic environments as shown in Equation (18-19).

$$\mathrm{W_j} = \begin{cases} \mathrm{W_j^{O1}}, \mathrm{E_j^{O1}} < \mathrm{E_j}; \\ \quad \mathrm{W_j}, \text{else}, \end{cases} \tag{18}$$

Where, $\mathrm{W_j}$ is the current candidate solution, $\mathrm{W_j^{O1}}$is the updated solution after exploration, $\mathrm{E_j^{O1}}$ and $\mathrm{E_j}$ are the corresponding fitness values, and the update is accepted only if improvement is achieved.

$$\mathrm{W_j} = \begin{cases} \mathrm{W_j^{O2}}, & \text{if } \mathrm{E_j^{O2}} < \mathrm{E_j} \\ \mathrm{W_j}, & \text{otherwise} \end{cases} \tag{19}$$

Where, $\mathrm{W_j^{O2}}$is the refined solution after exploitation, and $\mathrm{E_j^{O2}}$is its fitness value in Equation (21). The solution is updated only if it improves the objective function. RPO-Based Optimization for Context-Aware Representation Learning as shown in Algorithm 1. The design, optimization, training parameters, and regularization for performance enhancement of the proposed model are all summarized in Table 1.

Algorithm 1: RPO-VAE-ATT

Input: W (high-dimensional spatiotemporal data), S (time window), M (RPO population)
Output: Y_pred (crop yield), Z (context-aware latent features)
Initialize VAE parameters (θ, ϕ), attention weights, and RPO candidates W_j
Step 1: Preprocessing
W_ norm = normalize(W)
W_ feat = PCA (W_ norm)
Step 2: VAE Encoding
μ, σ = Encoder_ϕ(W_feat)
ε ~ N(0,1)
*Z = μ + σ * ε*
Step 3: Self-Attention
f = score(Z)
b = softmax(f)
*Y_context = sum(b * Z)*
Step 4: Decoding
W_ hat = Decoder_ θ(Z)

Step 5: Loss (ELBO)
*Loss = E[log pθ(W|Z)] - β * KL(qϕ(Z|W)||p(Z))*
Step 6: RPO Optimization
for each iteration:
  update W_ j (exploration & exploitation)
  select best solution minimizing Loss
Step 7: Prediction
Y_ pred = model(Y_ context)
return Y_ pred, Z

Table 1: Hyperparameter Configuration of the Proposed RPO-VAE-ATT Model

Category	Parameter	Value
Input	Dimension (D)	784
	Batch Size (B)	64
	Epochs (E)	100
VAE-Encoder	Layers	3
	Units	512, 256, 128
	Activation	ReLU
Latent Space	Dimension (d)	32
	μ, σ	Learned
Attention	Type	Attention
	Heads (h)	4
	Dimension (d_a)	64
	Dropout	0.2
VAE- Decoder	Layers	3
	Units	128, 256, 512
	Activation	Sigmoid
Loss	Reconstruction	MSE
	KL Weight (β)	1.0
RPO	Population (N)	30
	(T)	50
	α, γ, δ	0.7, 0.3, 0.5

Result and Discussion

In this results section, context-aware learning is demonstrated to improve predictive performance and adaptability in dynamic agricultural scenarios, judge accuracy and robustness, and compare the presented model with the existing methods. The model was implemented in Python 3.9 using Jupyter/Colab on Windows 10 , with Tensor Flow/PyTorch, standard libraries, Intel i7 CPU, optional NVIDIA GPU, 8–16GB RAM, and 80:20 training split.

The associations between the input features utilized in the suggested RPO-VAE-ATT model is shown in the Figure 4. Such correlations facilitate good feature selection and dimensionality reduction, which facilitates better representation learning and increases the accuracy, robustness and the overall crop yield prediction performance of the model. The proposed RPO-VAE-ATT approach makes use of these correlations to improve the model's overall crop yield prediction performance, accuracy, and robustness.

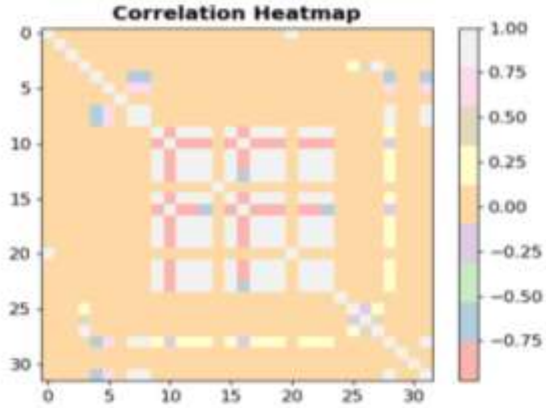

Figure 4: Correlation Heatmap of Feature Relationships in the Proposed RPO-VAE-ATT Framework

Figure 5(a) illustrates the distribution of crop output across multiple locations, highlighting variability, central tendency, and potential outliers, whereas Figure 5(b) demonstrates the changes in NDVI values over time, representing vegetation dynamics.These analyses illuminate time and location relationships of agricultural data. Through such patterns to construct context-aware representations, the proposed RPO-VAE-ATT method enhances feature elimination, dependency estimation, and crop yield forecast accuracy and robustness.

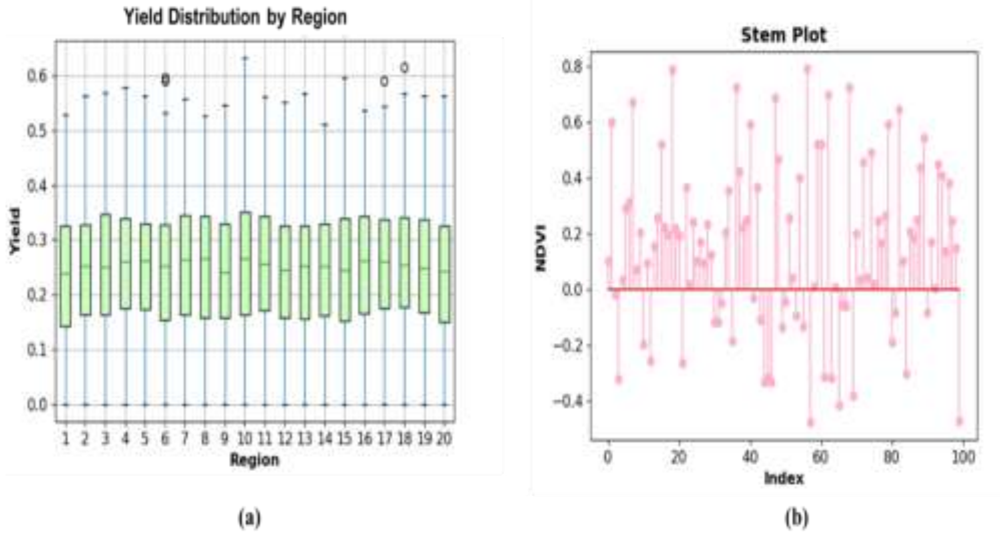

Figure 5: Visualization of Regional Yield Distribution and NDVI Variation (a) Regional Crop Yield Distribution Analysis, and (b) NDVI Temporal Variation

The allocation of NDVI values is shown in Figure 6 (a), which highlights density patterns and variability throughout the dataset. Figure 6 (b) shows the temporal dynamic relationship between NDVI and soil moisture. High-dimensional data feature comprehension is aided by these visual insights. The proposed RPO-VAE-ATT model incorporates the obtained patterns to construct context-sensitive representations, which increases the accuracy of the prediction and captures the spatial-temporal relationships in agricultural yield predictions efficiently.

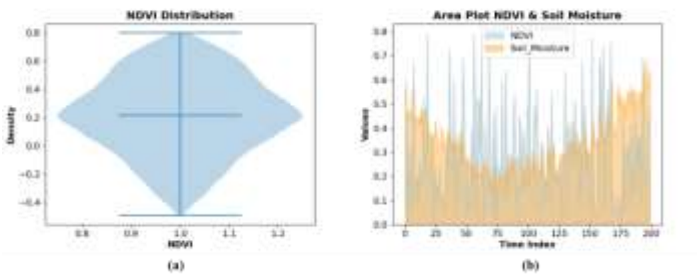

Figure 6: NDVI distribution and relationship with soil moisture across temporal data (a) NDVI Distribution Analysis, and (b) Temporal Variation of NDVI and Soil Moisture

Evaluation metrics

Accuracy represents the overall correctness of the proposed model in predicting crop yield outcomes based on geo-spatiotemporal agricultural data. Precision minimizes incorrect predictions by showing the percentage of accurately anticipated positive yield instances among all expected positive outcomes. Recall evaluates the framework's capacity to accurately detect instances of actual crop yield output without overlooking significant trends. F1-Score provides a fair assessment of the model's forecasting accuracy by calculating the harmonic average of precision and recall. Precision reflects the correctness of predicted positive outcomes, while Recall indicates the model's effectiveness in identifying all relevant instances.

Comparative Analysis of Context-Aware DL Models for Crop Yield Prediction

To compare the performance, the suggested framework was compared with Ensemble [19] and HDLM [20] under the same conditions. The findings indicate that the proposed model performed better with an accuracy of 95.3% rate, 96% rate, 97.39% rate, and 95.03% F1-score. Comparatively, Ensemble and HDLM had a relatively poorer performance in all metrics. The enhancement can be explained by the efficient preprocessing and learning of context-aware representations of spatial-temporal features. The general framework has proven to be more accurate and robust in predicting crop yield as indicated in Figure 7 and Table 2.

Table 2. Comparative Analysis of Crop Yield Prediction Model Performance Metrics

Model	Dataset	Accuracy (%)	F1-score (%)	Precision (%)	Recall (%)
Ensemble [19]	Rice Crop Yield Dataset	61	61	61	61
HDLM [20]	Custom dataset	90	94	92	96
RPO-VAE-ATT [Proposed]		95.3	95.03	96	97.39

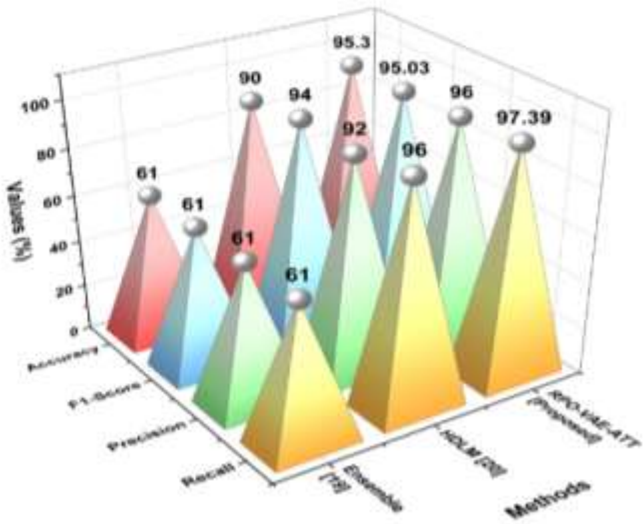

Fig 7. Comparison of model performance across metrics for different methods

DISCUSSION

The results emphasize that context-sensitive representation learning can well represent intricate spatial-temporal interactions in high-dimensional dynamic systems, which provide better predictive accuracy, stability and flexibility to predict accurate crop yields in different conditions. Despite its effectiveness, the method lacks context-aware representation learning to capture intricate spatial–temporal connections and is mostly dependent on climatological and distant sensing inputs, which limits adaptation to dynamic settings [15]. The model lacks the integration of multi-source data and sophisticated context-aware learning to handle high-dimensional dynamic agricultural systems, and despite enhanced explainability, it concentrates on early-season prediction and satellite-derived features [16]. To use remote sensing data to develop a fuzzy hybrid ensemble model to accurately classify crops and predict yield. Ensemble [19]: Computer complexity and training time: More complex than the other two approaches because it heavily relies on multiple base models. HDLM [20] was shown to perform poorly on unbalanced datasets and needed large-scale labelled data. The proposed RPO-VAE-ATT addresses the weaknesses of the earlier techniques with regard to context awareness and multi-source integration as it is able to discover dynamic spatial-temporal links to enhance prediction performance. The findings show that the suggested RPO-VAE-ATT framework uses context-aware representation learning to successfully capture intricate spatial–temporal connections. It is more resilient and offers better generalization than existing procedures due to the use of multi-source data and the best feature extraction.

Conclusion

Through efficiently learning context-aware representations from high-dimensional dynamic data, the proposed structure enhances crop yield forecast accuracy and robustness while facilitating trustworthy precision agriculture decision-making. The architecture combines multi-source geo-spatiotemporal data, which were pre-processed through Z-score normalization and PCA to extract features, and ensure high-quality inputs without any significant patterns of space and time alteration. A hybrid variational autoencoders with attention mechanism can learn the complex dependencies and semi-supervised learning can utilize the unlabeled data to improve performance in the limited labelled condition. The RPO-VAE-ATT model addresses the weaknesses of the current models such as the need of large datasets, single modality of data, and high computational needs, and provides reliable and generalized predictions. Experimental outcomes prove to work better with high accuracy (95.3%), precision (95.03%), recall (96%), and F1-score (97.39%), having a better robustness in various conditions. The framework's scalability and continuous deployment may be problematic, and it depends on the quality of data. Generalization is impacted by regional dataset limitations and a lack of labeled data. To improve predictive performance and flexibility, future research will concentrate on bigger, more varied datasets, current time implementation, system optimization, and interaction with IoT-based smart agriculture systems.

References

1. Nawaz, S., Saleem, M., Kusmartsev, F.V. and Anjum, D.H., 2024. Major role of multiscale entropy evolution in complex systems and data science. Entropy, 26(4), p.330.https://doi.org/10.3390/e26040330
2. Abdulrazzaq, M.M., Ramaha, N.T., Hameed, A.A., Salman, M., Yon, D.K., Fitriyani, N.L., Syafrudin, M. and Lee, S.W., 2024. Consequential advancements of self-supervised learning (SSL) in deep learning contexts. Mathematics, 12(5), p.758.https://doi.org/10.3390/math12050758
3. Rojas, L., Yepes, V. and Garcia, J., 2025. Complex dynamics and intelligent control: Advances, challenges, and applications in mining and industrial processes. Mathematics, 13(6), p.961.https://doi.org/10.3390/math13060961
4. Tomasetto, M., Williams, J.P., Braghin, F., Manzoni, A. and Kutz, J.N., 2025. Reduced order modeling with shallow recurrent decoder networks. Nature Communications, 16(1), p.10260.https://doi.org/10.1038/s41467-025-65126-y
5. EM, S., Chandy, D.A., PM, S. and Poulose, A., 2025. A Hybrid Deep Learning Model for Aromatic and Medicinal Plant Species Classification Using a Curated Leaf Image Dataset. AgriEngineering, 7(8), p.243. https://doi.org/10.3390/agriengineering7080243
6. Allam, H., Makubvure, L., Gyamfi, B., Graham, K.N. and Akinwolere, K., 2025. Text classification: How machine learning is revolutionizing text categorization. Information, 16(2), p.130.https://doi.org/10.3390/info16020130
7. Bai, X., Zhou, J. and Wang, Z., 2025. HPC Cluster Task Prediction Based on Multimodal Temporal Networks with Hierarchical Attention Mechanism. Computers, 14(8), p.335.https://doi.org/10.3390/computers14080335
8. Elhosseny, M., Abdel-Salam, M. and El-Hasnony, I.M., 2025. Adaptive dynamic crayfish algorithm with multi-enhanced strategy for global high-dimensional optimization and real-engineering problems. Scientific Reports, 15(1), p.10656.https://doi.org/10.1038/s41598-024-81144-0

9. Ruan, S., Wan, Q., Chen, R., Hu, M., Guo, X. and Song, K., 2026. Context-Aware Feature Enhancement Network for Remote Sensing Image Semantic Segmentation. Remote Sensing, 18(4), p.543.https://doi.org/10.3390/rs18040543
10. Sandeep, A., Jayarathna, S., Sandaruwan, S., Samarappuli, V., Meedeniya, D. and Perera, C., 2026. Context-Aware Multi-Agent Architecture for Wildfire Insights. Sensors, 26(3), p.1070.https://doi.org/10.3390/s26031070
11. Bakır, R. and Bakir, H., 2026. Enhancing SQL Injection Detection Using Contextual Natural Language Processing method and Artificial Intelligence Techniques. SN Computer Science, 7(4), p.291. https://doi.org/10.1007/s42979-026-04880-2
12. Kostopoulos, G., Stefani, A., Vasiliadis, V. and Kotsiantis, S., 2026. Deep Learning for e-Commerce: Recent Developments in Prediction, Personalization and Decision Intelligence. Applied Sciences, 16(5), p.2263.https://doi.org/10.3390/app16052263
13. Ponnumani, R., Vasudeva, N., Elumalai, T., Kaliyaperumal, P., Balusamy, B. and Benedetto, F., 2026. A multi-stage framework for scalable and context-aware intrusion detection in IoT-cloud systems using deep latent modeling and graph-based attack classification. Computers and Electrical Engineering, 131, p.110949.https://doi.org/10.1016/j.compeleceng.2026.110949
14. Hao, S., Yao, J., Shi, C., Zhou, Y., Xu, S., Li, D. and Cheng, Y., 2023. Enhanced semantic representation learning for sarcasm detection by integrating context-aware attention and fusion network. Entropy, 25(6), p.878.https://doi.org/10.3390/e25060878
15. Haseeb, M., Tahir, Z., Mahmood, S.A. and Tariq, A., 2025. Winter wheat yield prediction using linear and nonlinear machine learning algorithms based on climatological and remote sensing data. *Information Processing in Agriculture*. 12(4), p.431-444. https://doi.org/10.1016/j.inpa.2025.02.004
16. Fuentes, I. and Al-Shammari, D., 2026. Field-Aware and Explainable Modelling for Early-Season Crop Yield Prediction Using Satellite-Derived Phenology. *Remote Sensing*, *18*(6), p.890. https://doi.org/10.3390/rs18060890
17. Han, Y., Liu, X., Zhang, Y., Ai, H., Qin, C. and Zhang, X., 2026. FCR-TransUNet: A Novel Approach to Crop Classification in Remote Sensing Images Employing Attention and Feature Enhancement Techniques. *Agriculture*, *16*(7), p.727. https://doi.org/10.3390/agriculture16070727
18. Wei, C., Shan, Y. and Zhen, M., 2025. Deep learning-based anomaly detection for precision field crop protection. *Frontiers in Plant Science*, *16*, p.1576756. https://doi.org/10.3389/fpls.2025.1576756
19. Ilyas, Q.M., Ahmad, M. and Mehmood, A., 2023. Automated estimation of crop yield using artificial intelligence and remote sensing technologies. *Bioengineering*, *10*(2), p.125. https://doi.org/10.3390/bioengineering10020125
20. Chang, C.H., Lin, J., Chang, J.W., Huang, Y.S., Lai, M.H. and Chang, Y.J., 2024. Hybrid deep neural networks with multi-tasking for rice yield prediction using remote sensing data. *Agriculture*, *14*(4), p.513. https://doi.org/10.3390/agriculture14040513